



Universidade de Brasília  
Instituto de Ciências Exatas  
Departamento de Estatística

**MODELOS LINEARES MISTOS APLICADOS  
AOS DADOS DO PROGRAMA DOCE  
DESAFIO**

por

Ana Clara Ferreira Piantino e Fellipe da Silva Marques

Brasília

**2013**

**Ana Clara Ferreira Piantino e Fellipe da Silva  
Marques**

**MODELOS LINEARES MISTOS APLICADOS  
AOS DADOS DO PROGRAMA DOCE  
DESAFIO**

Relatório apresentado à disciplina Estágio Supervisionado II do curso de graduação em Estatística, Departamento de Estatística, Instituto de Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Professor Orientador: Eduardo Monteiro de Castro Gomes

Brasília

**2013**

Dedicamos este trabalho aos nossos pais e à Jane Dullius.

*Ana Clara Ferreira Piantino e Fellipe da Silva Marques*

# Agradecimentos

Agradecemos ao Prof. Eduardo Monteiro pela orientação deste trabalho, acrescentando conhecimento e correções necessárias para o sucesso.

Aos nossos pais que contribuíram para a nossa formação, além de incentivar incessantemente o nosso crescimento em todos os aspectos de nossas vidas.

Aos pais do Fellipe, por proporcionar o acolhimento em sua residência para as reuniões de elaboração do trabalho.

Aos colegas de curso e amigos, pelo incentivo durante o curso, pela valorização das nossas escolhas pessoais e profissionais, além do apoio nos momentos difíceis na realização deste programa final.

À coordenadora do programa doce DESAFIO, Jane Dullius, por proporcionar o banco de dados e informações para a conclusão de resultados.

Ao gerente do Fellipe, Reinaldo Soares de Camargo, pela compreensão das ausências para a conclusão do trabalho.

À ESTAT Consultoria por todo o aprendizado, autoconhecimento e motivação com o curso que foram proporcionados a Ana Clara nestes anos de empresa. Além da formação profissional e pessoal fomentada com a participação do Movimento Empresa Júnior.

# Resumo

Este trabalho aborda a aplicação de modelos lineares mistos utilizados sobre um banco de dados com medidas repetidas. A escolha do modelo envolveu a identificação dos efeitos aleatórios, dos efeitos fixos e da estrutura de covariâncias, utilizando-se técnicas gráficas e analíticas. Usando o teste de Razão de Verossimilhança e dos Critérios de Informação de Akaike – AIC e de Schwarz – BIC, pode-se fazer escolhas diferentes sobre a estrutura de covariância desejada, o que pode influenciar na seleção dos parâmetros de efeitos fixos. O trabalho se baseou em um conjunto de dados resultantes de um programa existente na Faculdade de Educação Física – FEF – UnB, localizado no centro olímpico desta universidade. Foi utilizada a função `lme()` *software* livre R – versão 3.0.1 (R,2013), devido a sua acessibilidade.

**Palavras-chave:** modelos lineares mistos, medidas repetidas, estrutura de covariância, Informação de Akaike, *Software* R – versão 3.0.1 (R,2013).

# Abstract

This paper discusses an application of linear mixed models used on a database with repeated measures. The choice of the model involved the identification of random effects, the fixed effects and the covariance structure using graphical and analytical techniques. Using the Likelihood Ratio test and the Akaike Information Criteria - AIC and Schwarz - BIC, we can make different choices about the desired covariance structure, which may influence the selection of the parameters fixed effects. This paper was based on a data set derived from an existing project in the Faculdade de Educação Física - FEF - UNB, located in the Olympic center of this university. We used the function `lme()` free *software* R – version 3.0.1 (R,2013), due to its accessibility.

**Keywords:** linear mixed models, repeated measures covariance structure, information Akaike, R. Software – version 3.0.1 (R,2013)

# Sumário

<b>RESUMO</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Objetivo</b>	<b>5</b>
<b>3 Metodologia</b>	<b>6</b>
3.1 Programa Doce DESAFIO . . . . .	6
3.2 Medidas Descritivas . . . . .	7
3.3 Introdução de Modelos Lineares . . . . .	8
3.4 Análise de Medidas Repetidas . . . . .	10
3.5 Modelos Lineares Mistos . . . . .	10
3.5.1 Modelo Linear Misto . . . . .	10
3.6 Estrutura das Matrizes de Covariância . . . . .	12
3.7 Método de Verossimilhança . . . . .	14
3.7.1 Máxima Verossimilhança Residual (MVR) . . . . .	15
3.8 Escolha do Modelo . . . . .	15
3.8.1 Critérios de Seleção . . . . .	16

3.9	Testes de Hipóteses de Seleção do Modelo . . . . .	17
3.9.1	Teste da Razão de Máxima Verossimilhança . . . . .	17
3.10	Análise de Resíduos . . . . .	17
<b>4</b>	<b>Análise</b>	<b>19</b>
4.1	Comparação entre as variáveis . . . . .	19
4.1.1	Análise de sexo por tipo de Diabetes . . . . .	20
4.1.2	Análise de sexo por insulino-terapia . . . . .	21
4.1.3	Análise de sexo por idade . . . . .	23
4.1.4	Análise de idade e insulino-terapia . . . . .	25
4.1.5	Análise de tipo de Diabetes e idade . . . . .	28
4.1.6	Análise de tipo de Diabetes e insulino-terapia . . . . .	29
4.1.7	Análise das glicemias iniciais e finais . . . . .	30
4.1.8	Glicemia iniciais e finais por insulino-terapia . . . . .	33
4.1.9	Análise de sexo e glicemia . . . . .	35
	Variação da glicemia . . . . .	35
	Glicemia inicial . . . . .	38
	Glicemia final . . . . .	38
4.1.10	Análise de glicemia por idade . . . . .	40
4.1.11	Análise de tipo de Diabetes e glicemias . . . . .	40
	Glicemia inicial . . . . .	41
	Glicemia final . . . . .	42
	Variação da glicemia . . . . .	42



4.1.12	Análise de tempo e glicemia . . . . .	43
4.2	Resultados da comparação com relação as aulas . . . . .	44
4.2.1	Comparação entre as aulas por sexo . . . . .	45
	Comparação entre glicemias finais, da primeira e última aula, do sexo masculino. . . . .	45
	Comparação das glicemias iniciais, da primeira aula, entre os sexos . . . . .	45
4.2.2	Comparação, entre as aulas, dos usuários de insulina exógena .	46
	Glicemia inicial na primeira e na última aula . . . . .	46
	Comparação das glicemias finais, na primeira e na última aula	47
4.2.3	Comparação das glicemias iniciais e finais das aulas, dos usuários e não usuários de insulina exógena . . . . .	47
	Comparação das glicemias iniciais, na primeira aula . . . . .	47
	Comparação das glicemias finais, na primeira aula . . . . .	48
	Comparação das glicemias finais, na última aula . . . . .	49
4.2.4	Comparação das médias iniciais e finais, por aula . . . . .	50
<b>5</b>	<b>Modelo dos dados</b>	<b>51</b>
5.1	Ajuste do modelo . . . . .	51
5.1.1	Modelos estimados . . . . .	52
	ANOVA de comparação entre os modelos com as variáveis significativas . . . . .	54

Comparação entre estruturas de covariância aplicadas ao mo-	
delo estimado . . . . .	55
Gráficos de resíduos . . . . .	55
<b>6 Conclusão</b>	<b>57</b>

# Lista de Tabelas

4.1	Tabela de Intervalo de confiança para as médias das glicemias . . . . .	20
4.2	Tabela de frequências de sexo por tipo de Diabetes . . . . .	20
4.3	Qui-Quadrado tipo de Diabetes e sexo . . . . .	21
4.4	Tabela de frequências de sexo por insulino-terapia . . . . .	21
4.5	Qui-Quadrado insulino-terapia e sexo . . . . .	22
4.6	Tabela de frequências de sexo por idade . . . . .	23
4.7	Teste-t de sexo por idade . . . . .	24
4.8	Tabela de frequências de insulino-terapia por idade . . . . .	25
4.9	Teste-t de insulino-terapia por idade . . . . .	27
4.10	Tabela de tipo de Diabetes por idade . . . . .	28
4.11	Teste-t de tipo de Diabetes por idade . . . . .	29
4.12	Tabela de tipo de Diabetes por insulino-terapia . . . . .	30
4.13	Teste-t de glicemia inicial e final . . . . .	33
4.14	Tabela de sexo e valores de variação da glicemia . . . . .	35
4.15	Teste-t de variação da glicemia por sexo . . . . .	36
4.16	Tabela de sexo e glicemias iniciais . . . . .	38
4.17	Tabela de sexo e valores de glicemias finais . . . . .	38

4.18	Tabela de tipo de Diabetes e glicemia inicial . . . . .	41
4.19	Teste-t de tipo de Diabetes e glicemia inicial . . . . .	41
4.20	Tabela de tipo de Diabetes e glicemia final . . . . .	42
4.21	Teste-t de tipo de Diabetes e glicemia final . . . . .	42
4.22	Tabela de tipo de Diabetes e valores de variação da glicemia . . . . .	42
4.23	Teste-t de tipo de Diabetes e variação da glicemia . . . . .	43
4.24	Correlações de Pearson entre as variáveis . . . . .	43
4.25	Teste-t entre as glicemias finais na 1ª e na última aula para o sexo masculino . . . . .	45
4.26	Teste-t das glicemias iniciais na 1ª aula entre os sexos . . . . .	46
4.27	Teste-t glicemia inicial dos usuários de insulina exógena entre as aulas	46
4.28	Teste-t glicemia final entre aulas . . . . .	47
4.29	Teste-t glicemia inicial primeira aula . . . . .	48
4.30	Teste-t glicemia final primeira aula . . . . .	48
4.31	Teste-t glicemia final última aula . . . . .	49
5.1	Modelo 1 (g1) : variação da glicemia $\sim$ tempo + sexo + insulino- terapia + idade + tipo de Diabetes . . . . .	52
5.2	Modelo 2 (g2) : variação da glicemia $\sim$ insulinoterapia + sexo + idade + tipo de Diabetes . . . . .	52
5.3	Modelo 3 (g3) : variação da glicemia $\sim$ insulinoterapia + idade + tipo de Diabetes . . . . .	53
5.4	Modelo 4 (g4) : variação da glicemia $\sim$ idade + tipo de Diabetes . .	53

5.5	Modelo 5 (g5) : variação da glicemia $\sim$ tipo de Diabetes . . . . .	53
5.6	Tabela de comparação entre os modelos estimados pelo método MV .	54
5.7	Tabela de comparação entre os métodos de estimação dos modelos . .	55

# Lista de Figuras

4.1	Gráfico tipo de Diabetes e sexo . . . . .	20
4.2	Gráfico insulino terapia e sexo . . . . .	22
4.3	Gráfico idade e sexo . . . . .	23
4.4	Boxplot insulino terapia e idade . . . . .	25
4.5	Boxplot tipo de Diabetes e idade . . . . .	29
4.6	Gráfico tipo de Diabetes e insulino terapia . . . . .	30
4.7	Boxplot de glicemia inicial e final . . . . .	31
4.8	Gráfico de glicemia inicial e final . . . . .	31
4.9	Boxplot de glicemia inicial e final dos insulino terapêuticos . . . . .	33
4.10	Boxplot variação da glicemia por sexo . . . . .	37
4.11	Boxplot glicemia inicial e final e sexo . . . . .	39
4.12	Diagrama de dispersão variação da glicemia e idade . . . . .	40
4.13	Gráficos de dispersão tempo e glicemias . . . . .	44
4.14	Médias das glicemias iniciais e finais por aula . . . . .	50
5.1	Gráfico de perfil dos participantes . . . . .	51
5.2	Gráfico dos resíduos padronizados . . . . .	55
5.3	Gráfico dos resíduos normalizados . . . . .	56

# Capítulo 1

## Introdução

Diabetes *Mellitus* é uma doença metabólica de origem múltipla, decorrente da falta de insulina e/ou da incapacidade de a insulina exercer adequadamente seus efeitos. A insulina é produzida pelo pâncreas e é responsável pela manutenção do metabolismo da glicose. A falta deste hormônio provoca déficit na metabolização da glicose e, conseqüentemente, o Diabetes. Caracteriza-se por altas taxas de açúcar no sangue (hiperglicemia) de forma permanente. Segundo Dullius (2007), a glicose é a principal fonte de energia do organismo porém, quando em excesso, pode trazer diversas complicações à saúde como, por exemplo: o excesso de sono no estágio inicial; problemas de cansaço e problemas físicos-táticos para efetuar as tarefas desejadas. Quando não tratada adequadamente, podem ocorrer complicações tais como: ataque cardíaco; derrame cerebral; insuficiência renal; problemas na visão; amputação do pé e lesões de difícil cicatrização; dentre outras complicações.

Embora ainda não haja a cura definitiva para a/o Diabetes (a palavra tanto pode ser feminina como masculina), há vários tratamentos disponíveis que, quando seguidos de forma regular, proporcionam saúde e qualidade de vida para o paciente. Diabetes é uma doença bastante comum no mundo, especialmente na América do

Norte e norte da Europa, acometendo cerca de 8% da população adulta entre 30 e 69 anos e 0,3% das gestantes no Brasil. Porém, estima-se que cerca de 50% dos portadores de Diabetes desconhecem o diagnóstico. Segundo projeção internacional, com o aumento do sedentarismo, obesidade e envelhecimento da população o número de pessoas com Diabetes no mundo vai aumentar em mais de 50%, passando de 380 milhões em 2025. (CAITANO P. et al., 2013)

Existem diversos tipos de Diabetes e, neste trabalho, foram evidenciados, o TIPO 1 (dependente de insulina exógena), é uma doença autoimune: as células do sistema imunológico atacam as células produtoras de insulina destruindo-as, resultando na diminuição ou cessação da produção de insulina. As pessoas com este tipo de Diabetes precisam aplicar injeções diárias de insulina, para controlar a doença que, em geral, inicia na infância ou na adolescência. Além deste, existe o TIPO 2, que corresponde a 90% dos diabéticos, segundo Dullius (2007), ocorre na grande maioria dos casos entre os adultos. Neste tipo de Diabetes a produção de insulina pelo pâncreas é reduzida e os tecidos do corpo se tornam resistentes à ação da insulina, impedindo a absorção da glicose pelo organismo, elevando, assim, a taxa de açúcar na corrente sanguínea.

Educação em Diabetes é o principal aspecto do tratamento, pois pouco adianta ter tecnologias e medicamentos avançados se o portador não souber como se cuidar e lidar com todas as adversidades que podem ocorrer em seu dia-a-dia. Um dos pilares do tratamento é a prática orientada de exercícios, que deve ser articulada com os demais aspectos da terapêutica tais como: nutrição; autocuidados; medicamentos;



acompanhamento psicossocial e exames clínicos, todos gerenciados pela Educação. (DULLIUS(2007))

A *American Diabetes Association*, segundo Zimonem B. e Ruderman N., afirma que a prática de exercícios físicos regulares é uma forma eficaz de ajuda no controle do Diabetes. Os principais benefícios do exercício em um indivíduo com Diabetes *Mellitus* TIPO 2 são: melhora a utilização de açúcar (glicose) pelos músculos devido ao aumento de gasto de energia no momento do exercício e a melhora da sensibilidade das células para a insulina, exatamente o que está com produção diminuída no Diabetes. Reduzindo a gordura corporal, tem-se melhor controle do Diabetes, o que auxilia na prevenção da doença e, a utilização da insulina pelas células, torna-se mais eficaz. Além disso, destaca-se a melhora da capacidade cardiorrespiratória, a circulação do corpo como um todo se beneficia, conseqüentemente, a ação de medicamentos utilizados no tratamento do Diabetes, torna-se melhor. Melhorando o controle do Diabetes fatores como motivação, autoestima e vontade para encarar novos desafios tornam-se mais presentes, diminuindo, por exemplo, a ansiedade, para comer que é um fator importante.

Os indivíduos que são fisicamente ativos têm menor chance de desenvolver Diabetes *Mellitus* TIPO 2, se comparados aos indivíduos sedentários.

Juntamente com essa linha de pensamento é realizado em Brasília um Programa chamado "Doce DESAFIO" – Programa Diabetes, Educação em Saúde e Atividades Físicas Orientadas. Neste programa, participam gratuitamente, diabéticos de todos os tipos (TIPO 1 e TIPO 2), idades e condições. Os participantes interagem em

grupos, por meio de debates sobre Diabetes e assuntos relacionados, situações do dia-a-dia. O objetivo é alertar os portadores da Diabetes que a melhor forma de ter uma vida saudável é aliar bons hábitos alimentares a exercícios físicos regulares. O programa mostra aos participantes o que se pode fazer para controlar a doença e aumentar cada vez mais a sua qualidade de vida.

# Capítulo 2

## Objetivo

O exercício físico é aceito como agente preventivo e terapêutico de diversas enfermidades. No tratamento de doenças cardiovasculares e crônicas, a atividade física tem sido apontada como principal medida não farmacológica, assumindo aspecto benéfico e protetor.

A melhora na saúde do indivíduo que pratica exercício físico regularmente é devido à exposição repetida do organismo a uma situação que requer uma reação mais forte, quando comparado com a sua atividade orgânica normal.

Levando em consideração a importância da atividade física, focada aos portadores de Diabetes, o objetivo desse trabalho é analisar o banco de dados do Programa "Doce DESAFIO", que tem como componente principal a medição da glicemia antes e depois da atividade física, no período de 24 aulas que compõem um semestre.

O intuito é verificar a variação da glicemia com a ajuda do exercício físico e dessa forma explicar, aplicar e simplificar, as informações para que os participantes do programa consigam entender de forma mais clara o que está acontecendo com o próprio corpo durante o programa e, os resultados positivos que são oferecidos pela atividade física.

# Capítulo 3

## Metodologia

### 3.1 Programa Doce DESAFIO

O Programa "Doce DESAFIO" abrange, e é a ampliação, do Proafidi (Programa Orientado de Atividades Físicas para Diabéticos). É um programa de extensão de ação contínua da faculdade de educação física da Universidade de Brasília, administrado pelo Decanato de Extensão da UnB como programa gratuito de educação em Diabetes, utilizando a atividade física como meio.

Direcionado aos Portadores de Diabetes *Mellitus* (PDMs) e dirigido por profissionais multidisciplinares especialmente qualificados da Universidade de Brasília, o Proafidi auxilia o tratamento de PDMs com a prática de atividades físicas, interação com os participantes favorecendo, também, a conscientização, a orientação, o acompanhamento e a atualização dos mesmos. Também, agrega conhecimento na formação de profissionais de educação física, e de outras áreas, no sentido de prover ambiente apropriado para o aprendizado e a experiência prática dos discentes.

O programa funciona principalmente no Centro Olímpico da UnB, Paranoá e Ceilândia mas são realizadas aulas esporádicas no Parque da Cidade. Já foram realizadas aulas , também, em: Sobradinho e Samambaia. Compõe-se de ativida-

des desenvolvidas duas ou três vezes por semana onde são observados: as idades, as condições clínicas, as capacidades e os interesses dos participantes. Em cada encontro são realizados anamnese (alimentação, medicação, condições gerais, observações), aferições glicêmicas e da pressão arterial, frequência cardíaca e outras quando necessário. São realizadas atividades físicas variadas conforme o grupo e as condições, entre elas: ginástica; jogos de bola; caminhadas; alongamento; natação; dança; corridas; hidroginástica; musculação; esportes diversos, etc.

Todas as informações são anotadas e mantidas em fichas individualizadas. (DULLIUS, 2007)

## **3.2 Medidas Descritivas**

A partir das anotações coletadas pelos colaboradores, de acordo com o objetivo deste trabalho, realizou-se a análise das variáveis disponíveis, tanto quantitativas quanto qualitativas. O banco de dados possui variáveis qualitativas como: aula, sexo, tipo de Diabetes (TIPO 1 e TIPO 2) e uso, ou não, de insulina exógena (insulinoterapia). As variáveis quantitativas são as seguintes: idade do aluno; tempo a partir da primeira aula; glicemia capilar antes e após a atividade física; além de uma variável que foi gerada, a variação da glicemia capilar.

A variação da glicemia capilar é calculada considerando a diferença entre a glicemia capilar final e glicemia capilar inicial. Essa medida foi calculada para cada um dos participantes em cada aula.

Os métodos iniciais utilizados para análise foram:

**análise descritiva:**

- médias das variáveis quantitativas;
- variâncias das variáveis quantitativas;
- valores máximos e mínimos;
- frequência observada das variáveis;
- gráficos para melhor visualização das frequências e distribuições das variáveis;

### **Inferências**

- correlação de Pearson
- testes de hipóteses:
  - teste *Qui-Quadrado* de Independência;
  - teste T- *Student*;
  - teste F de Fisher;

Com as análises dos dados, o propósito é: entender melhor as características dos participantes e a ação do programa em seus organismos e, realizar representações gráficas para aproximar os resultados dos participantes, com isso, são mostrados as alterações e avanços que a prática de exercício regular pode trazer ao controle do Diabetes e a sua qualidade de vida.

## **3.3 Introdução de Modelos Lineares**

Além das análises descritivas, o intuito do trabalho é encontrar um modelo que descreva o comportamento da glicemia capilar em pacientes com Diabetes *Mellitus*,

mais especificamente a variação da glicemia capilar. Avaliou-se quais variáveis eram significativas para o modelo, ou seja, quais variáveis influenciavam, significativamente, a variação da glicemia para esse grupo de pessoas.

Para orientar a proposição de algum modelo adequado para os dados em análise, é importante que se faça um estudo prévio acerca dos modelos mistos em análise estatística.

Os modelos lineares mais conhecidos em estatística são os modelos utilizados em estudos de regressão e delineamento de experimentos. Um modelo de regressão é utilizado quando se deseja caracterizar a relação entre uma variável dependente (variável resposta) e uma ou mais variáveis independentes (variáveis explicativas). Além de testar o efeito interativo de duas ou mais variáveis independentes, comparar múltiplos relacionamentos derivados da análise de regressão, entre outros.

A regressão linear simples tem como intuito descrever, a partir de uma relação matemática linear, o relacionamento entre duas variáveis. Entre todas as formas de utilização de regressão, destacam-se as seguintes:

- a partir de valores conhecidos de uma variável, estimar os valores de uma outra;
- confirmar uma relação de causa e efeito entre duas variáveis;
- previsão de valores futuros de uma variável.

Dados longitudinais são dados na forma de medições repetidas, na mesma unidade, ao longo do tempo. Segundo Lima (1996) para analisar os dados de um experimento com medidas repetidas existem técnicas utilizadas que vão desde análise de

variância uni e multivariada até a metodologia baseada em modelos lineares mistos, com modelagem da estrutura da matriz de variâncias e covariâncias.

### **3.4 Análise de Medidas Repetidas**

Segundo Diggle (1988) e também os autores Crowder e Hand (1990), dentre outros, o termo medidas repetidas, refere-se àqueles casos em que se observa uma ou mais variáveis resposta repetidamente na mesma unidade experimental.

Nesses estudos, as variáveis respostas podem ser contínuas (ganho de peso, consumo, conversão alimentar...). As unidades experimentais como indivíduos, plantas, animais, etc., podem estar classificadas em diferentes grupos, segundo um ou mais fatores (ou tratamentos), como sexo, tipo de Diabetes, uso de insulina exógena e etc.

### **3.5 Modelos Lineares Mistos**

Na Estatística, um tema que vem sendo crescente nas últimas décadas é modelos lineares mistos, na maioria das vezes presentes em dados agrupados, com a possibilidade de modelar a correlação intra-indivíduos. Observações repetidas sobre um mesmo indivíduo tendem a ser correlacionadas e os modelos lineares mistos são técnicas convenientes para modelar essa relação intra-indivíduos.

#### **3.5.1 Modelo Linear Misto**

Tratados primeiramente por Laird e Ware (1982), os modelos lineares mistos possibilitam o uso de várias estruturas de componentes de covariâncias. O primeiro passo é definir os efeitos fixos, para obtenção da curva polinomial média e, em seguida as



diferentes curvas para cada indivíduo.

Um modelo linear misto tem seguinte estrutura:

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i \quad (3.1)$$

Onde,  $\mathbf{Y}_i$  representa um vetor ( $n_i * 1$ ) de respostas da i-ésima unidade experimental ou indivíduo,  $\mathbf{X}_i$  é uma matriz ( $n_i * p$ ) de especificação (conhecida e de posto completo) dos efeitos fixos,  $\beta$  é um vetor ( $p * 1$ ) de parâmetros de (efeitos fixos),  $\mathbf{Z}_i$  é uma matriz ( $n_i * q$ ) de especificação (conhecida e de posto completo) dos efeitos aleatórios,  $\mathbf{b}_i$  é um vetor ( $q * 1$ ) de efeitos aleatórios com vetor de média 0 e matriz de covariância  $D$  e  $\epsilon_i$  é um vetor ( $n_i * 1$ ) com erros aleatórios, médias 0 e matriz de  $\text{cov}(\epsilon_i) = R_i(\boldsymbol{\lambda})$ , positiva definida, com  $\boldsymbol{\lambda}$  um vetor de parâmetros desconhecidos. Há  $N$  unidades experimentais e  $n_i$  observações feitas na i-ésima unidade experimental.

A forma de especificação da matriz  $\mathbf{X}_i$  é bastante similar àquela utilizada nos modelos de regressão e suas colunas podem estar relacionadas com:

- os fatores que definem a estrutura das subpopulações (grupos ou tratamentos);
- o tempo, definindo a curva a ser ajustada;
- as covariáveis, que os efeitos na resposta média desejam-se pesquisar;

Estas propriedades implicam que a matriz de covariância  $\text{cov}(\mathbf{Y}_i) = \Sigma = \mathbf{Z}_i D \mathbf{Z}_i^T + R_i$ . O modo estipulado anteriormente de estruturar a matriz de covariâncias  $\Sigma_i$  tem como atrativo a possibilidade de:

- abordar conceitos uni e multivariada que são frequentemente utilizadas na análise de dados longitudinais;
- lidar com dados perdidos, pois há a possibilidade de construir a Verossimilhança somente dos dados observados;
- abordar estruturas relacionadas com séries temporais ou estruturas mais complexas;

Quando ajustadas as curvas, os modelos lineares mistos assumem a existência de curvas subpopulacionais fixas  $(\mathbf{X}_i\beta)$ , com variações aleatórias  $(\mathbf{Z}_i\mathbf{b}_i)$  em tornos das curvas individuais, com as respectivas variações aleatórias  $(\epsilon_i)$  em torno da curva média. O modelo linear misto, normalmente, é utilizado em termos das respostas condicionadas aos efeitos aleatórios de modo que  $\mathbf{Y}_i$  assumido como um vetor de medidas repetidas para o  $i$ -ésimo indivíduo, satisfaça:

$$\mathbf{Y}_i|\mathbf{b}_i \sim N(\mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i, R_i) \quad (3.2)$$

$$\mathbf{b}_i \sim N(0, D) \quad (3.3)$$

### 3.6 Estrutura das Matrizes de Covariância

As estimativas e os erros padrões de efeitos fixos, diagnósticos e inferências são influenciados de acordo com a forma da estrutura de covariância selecionada, dependendo da informação empírica da estrutura de dados e, muitas vezes da possibilidade computacional.

A estrutura dos modelos lineares mistos possibilita a consideração de matrizes especiais de covariância, que buscam representar de forma eficaz e precisa a variabilidade dos dados, ou seja, consegue absorver as diferentes informações, levando em conta se os dados são independentes, dependentes, correlacionados, dentre outros. (BARBOSA, 2009)

A seguir, algumas estruturas da matriz  $R_i$  implementadas no *Software R*.

### 1. Matriz de componentes de variância (VC)

Variâncias iguais nas  $n_i$  observações e independentes entre elas:

$$\begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}$$

### 2. Matriz de simetria composta (CS)

Variâncias iguais nas  $n_i$  observações e mesma covariância entre medidas feitas em tempos diferentes.

$$\begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 \end{pmatrix}$$

### 3. Não Estruturada

Variâncias diferentes para cada uma das  $n_i$  ocasiões e covariâncias distintas entre medições feitas em ocasiões distintas.  $n_i(n_i+1)/2$  parâmetros:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{33} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{43} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

### 3.7 Método de Verossimilhança

Em modelos lineares mistos, além de novos métodos de refinamentos apresentados todos os dias, são utilizados basicamente dois métodos: Método de Verossimilhança e o Método de Verossimilhança Restrita.

Em Estatística, a estimativa por Máxima Verossimilhança (*Maximum-Likelihood Estimation- MLE*) é um método de estimação dos diferentes parâmetros de um modelo. A partir de um conjunto de dados pode ser encontrado um modelo estatístico.

Esses parâmetros da distribuição podem então ser estimados por MLE a partir da medição de uma amostra da população. O método busca valores para os parâmetros de maneira a maximizar a probabilidade dos dados amostrados, dado o modelo assumido (no caso, distribuição Normal).

Então, a inferência associada aos modelos lineares mistos é baseada na densidade marginal do vetor de resposta  $\mathbf{Y}_i$ . Dessa forma, Verbeke e Molenbergs (2000) apresentam a forma clássica de estimação baseada na maximização da função Verossimilhança Marginal:

$$L_V(\theta) = \prod_{i=1}^n \{(2\Pi)^{-n_i/2} |\Sigma_i(\alpha)|^{-1/2} \exp(-1/2(y_i - X_i\beta)'(\Sigma_i^{-1}(\alpha))(y_i - X_i\beta))\} \quad (3.4)$$

Onde,  $\alpha$  é o vetor com todos os parâmetros de variâncias e covariâncias encontrado  $\Sigma_i = Z_i D Z_i^T + R_i$ , que consiste  $n_i(n_i + 1)/2$  elementos diferentes em  $D$  e de todos os parâmetros em  $R_i$ . Seja  $\theta$  o vetor de parâmetros para o modelo marginal  $y_i$ , ou seja, o vetor de parâmetros da variável resposta (variação da glicemia).

O estimador de Verossimilhança de  $\beta$ , a partir da maximização da função, con-

dicionada a  $\alpha$  segundo Laird e Ware (1982), é dada por:

$$\hat{\beta}(\alpha) = \sum_{i=1}^N (X_i' \Sigma_i^{-1} X_i)^{-1} \sum_{i=1}^N (X_i' \Sigma_i^{-1} y_i) \quad (3.5)$$

### 3.7.1 Máxima Verossimilhança Residual (MVR)

O método de Máxima Verossimilhança Residual (MVR) proposto por Patterson e Thompson (1971), também pode ser chamado de Máxima Verossimilhança Restrita e, possui o objetivo de estimar os componentes da variância. A estimação de  $\sigma^2$  seria viesada se fosse utilizado o Método da Verossimilhança, pois estaria subestimando as estimativas dos parâmetros. O Método de MVR está relacionado aos contrastes dos erros e não às observações reais, fazendo assim com que as estimativas dos componentes da variância apresentem estimadores com menor viés, segundo Davis (2002). A estimação por Verossimilhança Restrita é dada pela seguinte equação:

$$L_{VR}(\theta) = \left| \sum_{i=1}^N (X_i' \Sigma_i^{-1} X_i) \right|^{-1/2} L_V(\theta) \quad (3.6)$$

Com a intenção de maximizar o logaritmo da Verossimilhança Restrita são necessários métodos iterativos, como método de Fisher Scoring e método de Newton-Raphson, sendo o primeiro considerado melhor em relação ao tempo total para atingir a convergência, a partir das modificações propostas por Laird e Ware (1982).

## 3.8 Escolha do Modelo

Selecionar o melhor modelo, significa verificar a melhor forma de estruturar a parte fixa (efeito fixo) e a covariância, mas não somente isso. É necessário investigar

e identificar os efeitos aleatórios, bem como a estimação e comparação entre modelos. Existem critérios na literatura para a escolha do melhor modelo que auxiliam a decisão.

Considerando os métodos para selecionar o modelo mais adequado, torna-se mais acessível encontrar um modelo que represente adequadamente os dados.

### 3.8.1 Critérios de Seleção

O Critério de Informação de Akaike (AIC) admite a existência de um modelo real que descreve os dados, que é desconhecido, e tenta escolher dentre um grupo de modelos a serem avaliados. A estimativa do AIC para um determinado modelo é dada por:

$$AIC = -2l + 2k \quad (3.7)$$

Onde,  $l$  é o máximo da log-verossimilhança e  $k$  o número de parâmetros. O modelo com menor valor de AIC é considerado o modelo de melhor ajuste.

O Critério Bayesiano de Schwarz (BIC) tem como pressuposto a existência de um modelo verdadeiro que descreve a relação entre a variável dependente e as diversas variáveis explicativas dos modelos. O valor do critério BIC para um determinado modelo é dado por:

$$BIC = -2l + k \log(n) \quad (3.8)$$

Com  $n$  o número de observações. O modelo com menor BIC é considerado o de melhor ajuste.

Ambos os critérios penalizam modelos com muitas variáveis sendo que valores menores de AIC e BIC são preferíveis.

## **3.9 Testes de Hipóteses de Seleção do Modelo**

### **3.9.1 Teste da Razão de Máxima Verossimilhança**

Para os modelos ajustados pelo método da Máxima Verossimilhança, o teste mais utilizado para seleção do melhor modelo é o Teste de Razão de Verossimilhança (PINHEIROS E BATES, 2000). O Teste é baseado na razão entre as Verossimilhanças dos dois modelos ajustados, para ser verificado se o modelo reduzido (menor número de parâmetros) se ajusta tão bem quanto o modelo completo (com a quantidade total do número de parâmetros).

A estatística do Teste é dada por:

$$\xi_{RV} = 2[\log(L_2) - \log(L_1)] \quad (3.9)$$

Com distribuição qui-quadrado com  $r$  graus de liberdade, onde  $r$  é a diferença entre o número de parâmetros dos modelos testados,  $L_1$  é o valor da função de Verossimilhança do modelo reduzido e  $L_2$  do modelo completo.

## **3.10 Análise de Resíduos**

A análise de resíduos é essencial no ajuste do modelo, independente de sua natureza, desde os modelos lineares mais simples aos modelos generalizados lineares e não lineares, como também os modelos mais complexos.

É importante dizer que cada modelo detém uma determinada estrutura, apesar

da metodologia de ajuste. Assim, o estudo da adequabilidade de forma geral tem a intenção de verificar as suposições do modelo e a forma como as observações influenciam no seu ajuste. Vale ressaltar, que é necessária uma abordagem cuidadosa da análise dos resíduos levando em consideração que a estrutura dos resíduos varia de modelo para modelo no estudo da adequabilidade.

Seguindo a referência (Pinheiro e Bates, 2000), propõe-se o uso de gráficos de resíduos condicionais em relação aos valores ajustados para avaliar a suposição de homocedasticidade e o uso do gráfico de probabilidade normal dos resíduos condicionais para avaliar a suposição de normalidade. Vale evidenciar que os gráficos de resíduos condicionais também podem ser utilizados para identificar pontos discrepantes.

Espera-se que os resíduos do modelo variem aleatoriamente em torno do zero, para suposição da aleatoriedade dos resíduos.



# Capítulo 4

## Análise

### 4.1 Comparação entre as variáveis

O Programa "Doce DESAFIO" coletou dados de 85 alunos que frequentaram 24 aulas consecutivas, um semestre, sendo essas as informações presentes no banco de dados que foi utilizado nesse trabalho. Realizou-se a análise descritiva das variáveis: sexo, tipo de Diabetes, uso de insulina exógena, idade, glicemia inicial, glicemia final, variação das glicemias e tempo a partir da primeira aula de cada participante, além da relação entre essas variáveis.

Os dados fornecidos apresentam: 53% de participantes do sexo masculino e 47% do sexo feminino. Como explicado na introdução deste trabalho, o Diabetes *Mellitus* apresentou dois tipos principais: TIPO 1 e 2. Os dados apresentam 16% dos indivíduos portadores de Diabetes TIPO 1 e 84% portadores de Diabetes TIPO 2. Dentre os participantes, 45% são usuários de insulina exógena e 55% não usuários de insulina exógena. Nos cálculos de intervalos de confiança com nível de significância ( $\alpha$ ) igual a 5% para as médias das glicemias iniciais, finais e variações das glicemias, obteve-se:

Tabela 4.1: Tabela de Intervalo de confiança para as médias das glicemias

Variável	Média	Intervalo de Confiança
Glicemia inicial	168,06	[155,31 ; 180,81]
Glicemia final	122,71	[113,48 ; 131,93]
variação das glicemias	-45,41	[-117,87 ; 27,05]

#### 4.1.1 Análise de sexo por tipo de Diabetes

A partir das informações coletadas, segue a tabela de frequências relacionada:

Tabela 4.2: Tabela de frequências de sexo por tipo de Diabetes

	TIPO		
<i>Sexo</i>	1	2	<i>Total</i>
Masculino	6	39	45
Feminino	8	32	40
Total	14	71	85

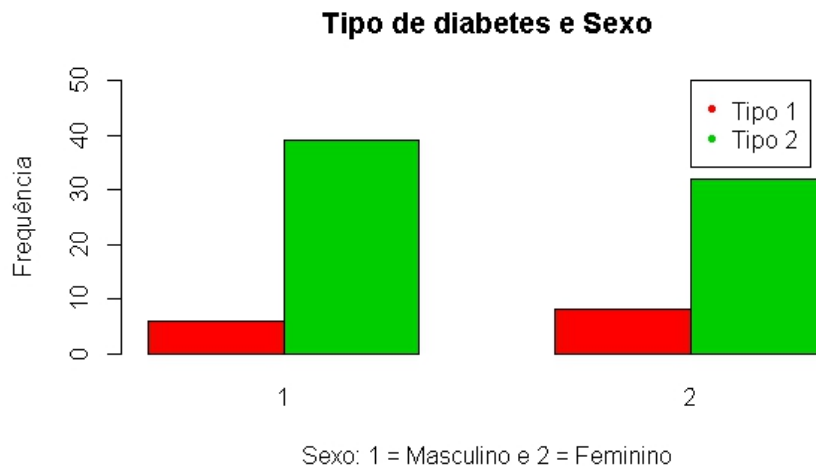


Figura 4.1: Gráfico tipo de Diabetes e sexo

Para avaliar a relação entre as variáveis, foi feito teste *Qui-Quadrado* de independência com as seguintes hipóteses:

- $H_0$ : o tipo de Diabetes e sexo são independentes.
- $H_a$ : o tipo de Diabetes e sexo não são independentes.

Tabela 4.3: Qui-Quadrado tipo de Diabetes e sexo

	$\alpha$	Graus de liberdade	$\chi^2$	p-valor
Qui-Quadrado	0,05	1	0,2853	0,5932

A partir do teste *Qui-Quadrado*, com um nível de significância de 5%, rejeita-se a hipótese nula se o p-valor for menor que  $\alpha$ . Neste caso, com p-valor igual a 0,5932, alto e maior que  $\alpha$ , não rejeita-se a hipótese nula que há independência entre as variáveis. Sendo assim, pode-se dizer que as variáveis sexo e tipo de Diabetes são independentes, ou seja, o sexo do participante não está diretamente ligado ao tipo de Diabetes. Vale acrescentar que o tipo de Diabetes mais presente no banco é o TIPO 2, assim como na população geral onde o Diabetes do TIPO 2 é mais comum, por ter uma possível relação com maus hábitos alimentares e de sedentarismo.

#### 4.1.2 Análise de sexo por insulino terapia

Tabela 4.4: Tabela de frequências de sexo por insulino terapia

	Insulino terapia		
Sexo	Sim	Não	Total
Masculino	24	21	45
Feminino	14	26	40
Total	38	47	85

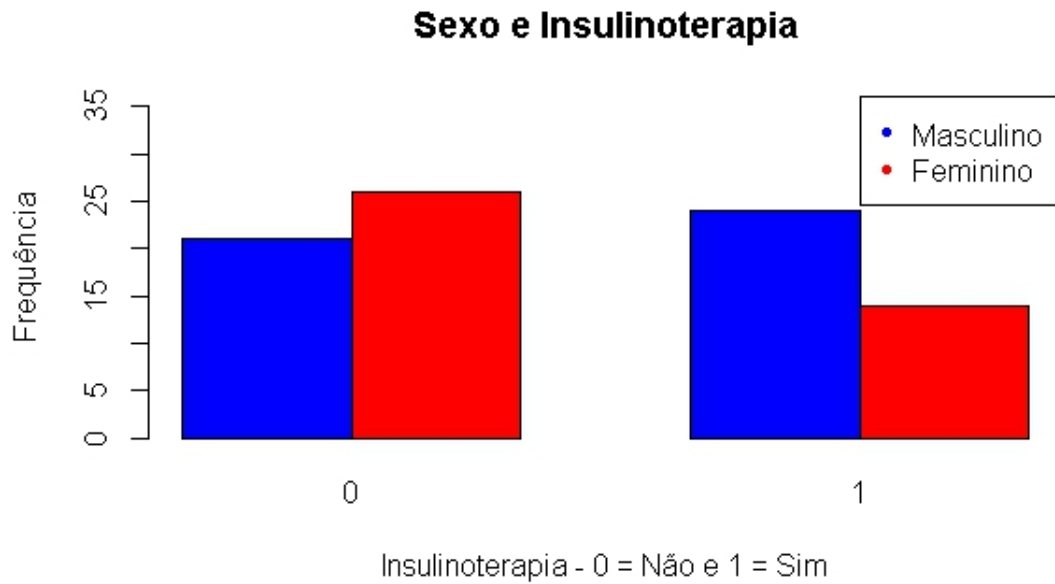


Figura 4.2: Gráfico insulinoterapia e sexo

Para avaliar a relação entre as variáveis, foi feito teste *Qui-Quadrado* de independência com as seguintes hipóteses:

- $H_0$ : a insulinoterapia e o sexo do participante são independentes.
- $H_a$ : a insulinoterapia e o sexo do participante não são independentes.

Tabela 4.5: Qui-Quadrado insulinoterapia e sexo

	$\alpha$	Graus de liberdade	$\chi^2$	p-valor
Qui-Quadrado	0,05	1	2,1854	0,1393

A partir do gráfico 4.2, pode-se perceber que a insulinoterapia é mais incidente nos homens que participam do programa. A necessidade de insulinoterapia é a única opção de tratamento do Diabetes do TIPO 1, já para o Diabetes TIPO 2 há mais opções, além da insulinoterapia. Por ser uma doença progressiva, conforme o envelhecimento, maior a chance da necessidade de reposição de insulina.

Analisando o teste *qui-quadrado* referente as variáveis sexo e insulinoterapia, encontra-se um p-valor maior que o nível de significância ( $\alpha$ ), levando a não rejeição da hipótese nula, ou seja, as variáveis são independentes.

### 4.1.3 Análise de sexo por idade

Tabela 4.6: Tabela de frequências de sexo por idade

<i>Idade</i>	Sexo		<i>Total</i>
	<i>Masculino</i>	<i>Feminino</i>	
0 -  15	2	5	7
15 -  30	5	1	6
30 -  45	0	2	2
45 -  60	20	17	37
60 -  75	18	15	33
Total	45	40	85

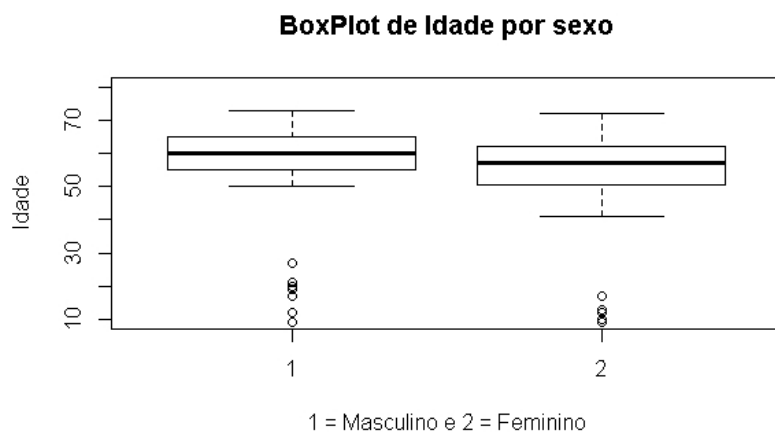


Figura 4.3: Gráfico idade e sexo

Analisando a idade dos participantes de ingresso no programa em relação ao sexo, realizou-se um teste-t para comparação de médias das idades.

Para realizar a comparação, analisou-se primeiramente se as variâncias das idades em cada sexo eram iguais, para isso foi feito um teste-F de Fisher para comparar essas variâncias, onde a hipótese nula é que as variâncias dos grupos são iguais.

O p-valor do teste foi igual a 0,5502, maior que o nível de significância de 5%, ou seja, não rejeita-se a hipótese nula. Não existe diferença significativa entre as variâncias das idades por sexo.

A partir dessas conclusões e sabendo que há diferença também entre a quantidade de pessoas de cada sexo, o teste mais adequado para comparar as médias de idade por sexo é o teste-t para amostras independentes com variâncias iguais.

As hipóteses do teste-t nesse caso, são as seguintes:

- $H_0$ : a média das idades do sexo feminino é igual a média das idades do sexo masculino.
- $H_a$ : a média das idades do sexo feminino é menor do que a média das idades do sexo masculino.

Tabela 4.7: Teste-t de sexo por idade

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	0,6954	0,2444

Observando as médias de idades por sexo para realizar o teste-t, foi encontrado um p-valor alto. Sendo assim, não se rejeita a hipótese nula. Logo, não há evidências para afirmar que existe diferença significativa entre as médias das idades entre os sexos.

Tabela 4.8: Tabela de frequências de insulinoaterapia por idade

Idade	Insulinoterapia		Total
	Não	Sim	
0 -  15	0	7	7
15 -  30	0	6	6
30 -  45	1	1	2
45 -  60	22	15	37
60 -  75	24	9	33
Total	47	38	85

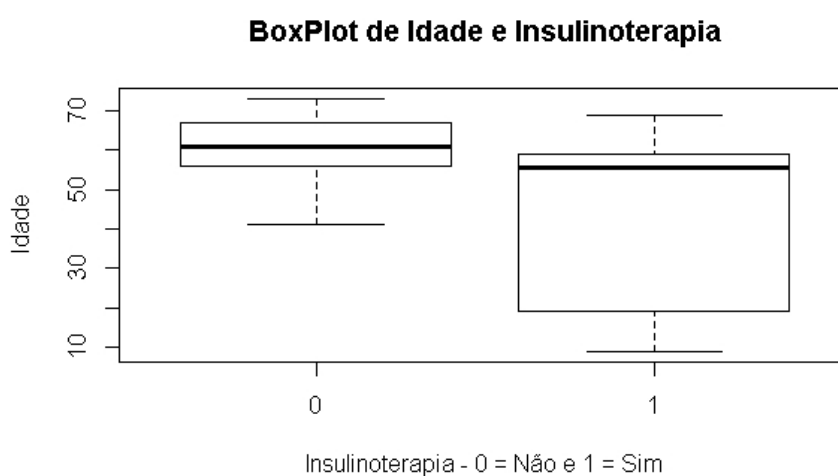


Figura 4.4: Boxplot insulinoterapia e idade

#### 4.1.4 Análise de idade e insulinoterapia

Comparando as variáveis insulinoterapia (0 = não utiliza insulina exógena e 1 = utiliza insulina exógena) e as médias das idades, observa-se pelo gráfico 4.4 que as médias de idades são diferentes com relação à necessidade de insulina exógena.

Essa diferença pode ser explicada pelo fato de a insulinoterapia ser o único tratamento para quem tem Diabetes do TIPO 1, que é o tipo mais frequente em pessoas com idade mais jovens. Analisando a tabela 4.9, percebe-se que todos os participan-

tes até os 30 anos são dependentes de insulina exógena, o que influencia nas médias de idade dos indivíduos que necessitam de insulino-terapia.

Observando o gráfico 4.4, nota-se que as idades de quem necessita de insulino-terapia tem maior variação, apresentada em todos os intervalos de idades enquanto os que não necessitam da reposição concentram-se a partir do intervalo de 30 a 45 anos. Evidencia-se a presença maior de insulino-terapia em idades precoces, mas há também, a presença de diabéticos TIPO 2 com insulino-terapia, por ser uma doença progressiva e ao longo do tempo a chance da necessidade do uso de insulina exógena é maior. Por isso, percebe-se ,também, presença de insulino-terapia nas idades mais avançadas.

Analisando a idade dos participantes, de ingresso no programa em relação a necessidade, ou não, de insulino-terapia, realizou-se um teste-t para comparação de médias das idades.

Para realizar a comparação, analisou-se primeiramente se as variâncias das idades nos dependentes, ou não, de insulina exógena eram iguais, para isso foi feito um teste-F de Fisher para comparar essas variâncias, onde a hipótese nula é que as variâncias dos grupos são iguais.

O p–valor do teste foi igual a  $6,92\epsilon^{-12}$ , menor que o nível de significância de 5%, ou seja, rejeita-se a hipótese nula. Existe diferença significativa entre as variâncias das idades por insulino-terapia.

A partir dessas conclusões e sabendo que há, também, diferença entre a quantidade de pessoas classificadas na variável insulino-terapia, o teste mais adequado



para comparar as médias de idade por insulino-terapia é o teste-t para amostras independentes com variâncias diferentes.

As hipóteses do teste-t nesse caso, são as seguintes:

- $H_0$ : a média das idades dos dependentes de insulina exógena é igual a média das idades dos não dependentes de insulina exógena.
- $H_a$ : a média das idades dos dependentes de insulina exógena é menor a média das idades dos não dependentes de insulina exógena.

Tabela 4.9: Teste-t de insulino-terapia por idade

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	5,25	$5,747\epsilon^{-7}$

Observando as médias de idades por insulino-terapia para realizar o teste-t, foi encontrado um p-valor baixo, menor que  $\alpha$ , sendo assim, rejeita-se a hipótese nula. Logo, há evidências para afirmar que as médias das idades dos participantes que necessitam da insulina exógena são menores que as médias das idades daqueles que não necessitam da reposição.

### 4.1.5 Análise de tipo de Diabetes e idade

Agora, comparando o tipo de Diabetes e a idade dos participantes, obtém-se:

Tabela 4.10: Tabela de tipo de Diabetes por idade

Tipo de Diabetes	Idade		
	Mínima	Média	Máxima
TIPO 1	9	18,35	48
TIPO 2	27	59,9	73

Analisando a idade dos participantes ao entrarem no programa em relação ao tipo de Diabetes, realizou-se testes para comparação das médias de idades e tipo de Diabetes.

Para realizar a comparação, analisou-se primeiramente se as variâncias das idades com relação ao tipo de Diabetes eram iguais. Para isso foi feito um teste-F de Fisher para comparar essas variâncias, onde a hipótese nula é que as variâncias dos grupos são iguais.

O p-valor do teste foi 0,011, menor que o nível de significância de 5%, ou seja, rejeita-se a hipótese nula. Existe diferença significativa entre as variâncias das idades e o tipo de Diabetes.

Levando em consideração que há diferença entre a quantidade de pessoas de cada tipo de Diabetes, o teste mais adequado para comparar as médias de idade por tipo de Diabetes é o teste-t para amostras independentes com variâncias diferentes.

As hipóteses do teste-t nesse caso, são as seguintes:

- $H_0$ : a média de idades dos diabéticos do TIPO 1 é igual a média de idades dos diabéticos do TIPO 2.

- $H_a$ : a média de idades dos diabéticos do TIPO 1 é menor que a média de idades dos diabéticos do TIPO 2.

Tabela 4.11: Teste-t de tipo de Diabetes por idade

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	-17,381	$2,2\epsilon^{-16}$

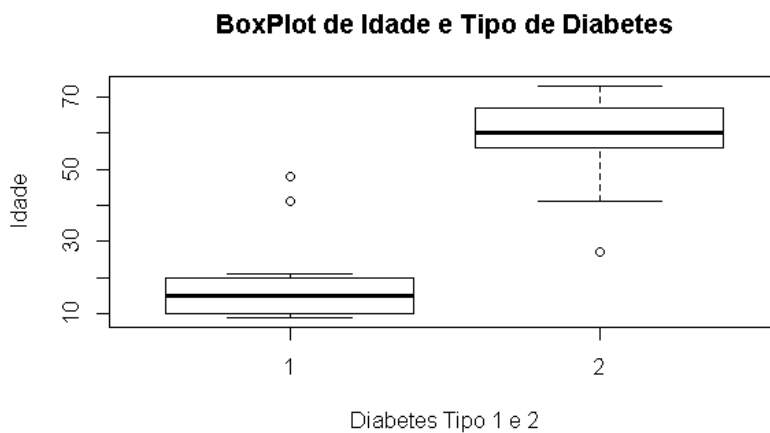


Figura 4.5: Boxplot tipo de Diabetes e idade

Analisando a relação entre idade e tipo de Diabetes, verifica-se que a média de idades para indivíduos portadores de Diabetes *Mellitus* TIPO 1 é menor que a média de idades dos portadores de Diabetes *Mellitus* TIPO 2. Confirmando-se assim que o Diabetes *Mellitus* de TIPO 2 é mais frequente em participantes com idades mais avançadas e a Diabetes *Mellitus* TIPO 1 mais frequente em adolescentes e crianças.

#### 4.1.6 Análise de tipo de Diabetes e insulinoterapia

Os resultados são apresentador a seguir:

Tabela 4.12: Tabela de tipo de Diabetes por insulino terapia

Tipo de Diabetes	Insulinoterapia		Total
	Não	Sim	
TIPO 1	0	14	14
TIPO 2	47	24	71
Total	47	38	85

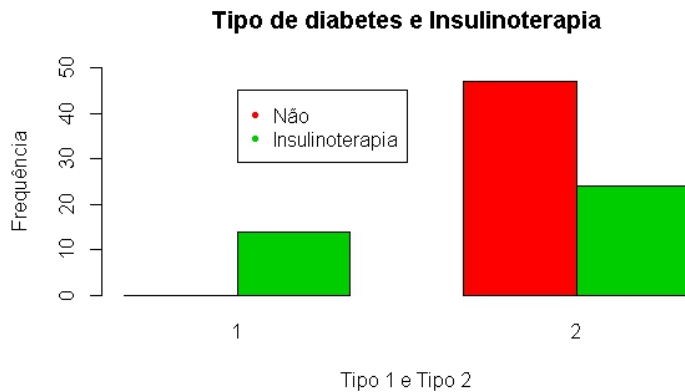


Figura 4.6: Gráfico tipo de Diabetes e insulino terapia

Por se tratarem de duas variáveis qualitativas, seria adequada a aplicação de um teste para independência. Na tabela 4.12, há restrições para a aplicação do teste. Um exemplo é a frequência do TIPO 1 e não uso de insulina exógena, onde a frequência é 0.

A partir das conclusões dos tópicos anteriores, observa-se que 100% dos diabéticos TIPO 1 e 33,8% dos diabéticos TIPO 2 necessitam do uso de insulina exógena.

#### 4.1.7 Análise das glicemias iniciais e finais

Condizente com o objetivo desse trabalho, foram analisadas as taxas de glicemia no início e no fim de cada uma das 24 aulas.

A partir de uma análise gráfica das variáveis disponíveis, observa-se no box-plot

4.7 a diferença clara de posição entre as glicemias iniciais e finais dos indivíduos, pode-se dizer que há uma tendência de redução das glicemias ao longo da prática do exercício físico, pois houve uma queda na glicemia média inicial para a glicemia média final.

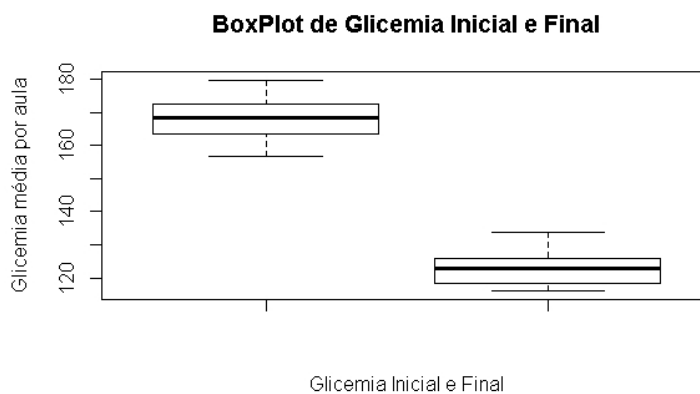


Figura 4.7: Boxplot de glicemia inicial e final

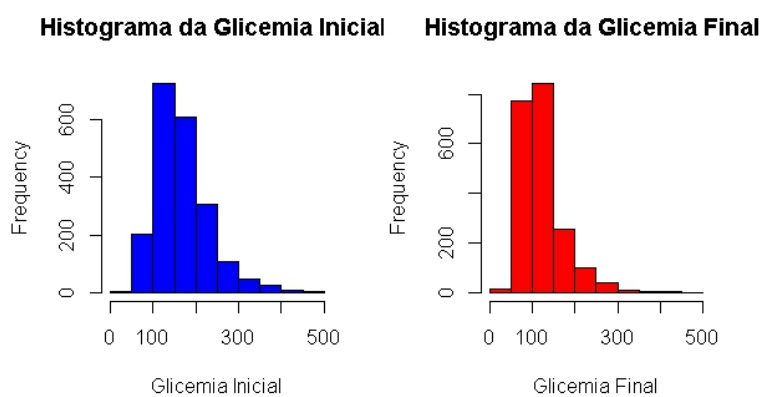


Figura 4.8: Gráfico de glicemia inicial e final

O histograma representado na figura 4.8 também mostra que as glicemias depois do exercício físico são mais baixas, além de mostrar que a distribuição dos dados não está tão próxima da distribuição Normal.

Afim de perceber a relação entre as glicemias de cada participante, foi calculado a média das glicemias por aluno, assim cada aluno obteve uma média de glicemia inicial referente a todas as aulas e o mesmo se aplica para a glicemia final e a variação.

Para testar o que foi apresentado no gráfico 4.8, realizou-se um teste-t para comparar as médias das glicemias antes e depois da atividade física.

Para realizar a comparação, analisou-se primeiramente se as variâncias das glicemias iniciais e finais eram iguais. Para isso foi realizado um teste-F de Fisher para comparar essas variâncias, onde a hipótese nula é que as variâncias dos grupos são iguais.

O p-valor do teste foi menor que 0,002, menor que o nível de significância de 5%, ou seja, rejeita-se a hipótese nula. Existe diferença significativa entre as variâncias das glicemias iniciais e finais.

Sendo assim, o teste mais adequado para comparar as médias das glicemias antes e depois do exercício físico é o teste-t para amostras pareadas com variâncias diferentes.

As hipóteses do teste-t nesse caso, são as seguintes:

- $H_0$ : a média da glicemia inicial é igual a média da glicemia final.
- $H_a$ : a média da glicemia inicial é maior a média da glicemia final.

A partir do teste-t, rejeita-se a hipótese nula com o nível de significância de 5%, ou

Tabela 4.13: Teste-t de glicemia inicial e final

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	168,00	8,29	$1,737\epsilon^{-14}$

seja, pode-se afirmar que a média da glicemia inicial (168,0647) é significativamente maior que a média da glicemia final(122,7157), confirmando a efetividade do exercício físico na redução da glicemia dos participantes.

#### 4.1.8 Glicemia iniciais e finais por insulinoterapia

Considerou-se as análises com relação aos participantes que tem a necessidade da insulinoterapia, onde foram apresentadas diferenças significativas entre os que necessitam, ou não, de insulinoterapia. Notou-se que seria interessante comparar as glicemias iniciais e finais comparando os participantes que usam insulina exógena e os que não usam.

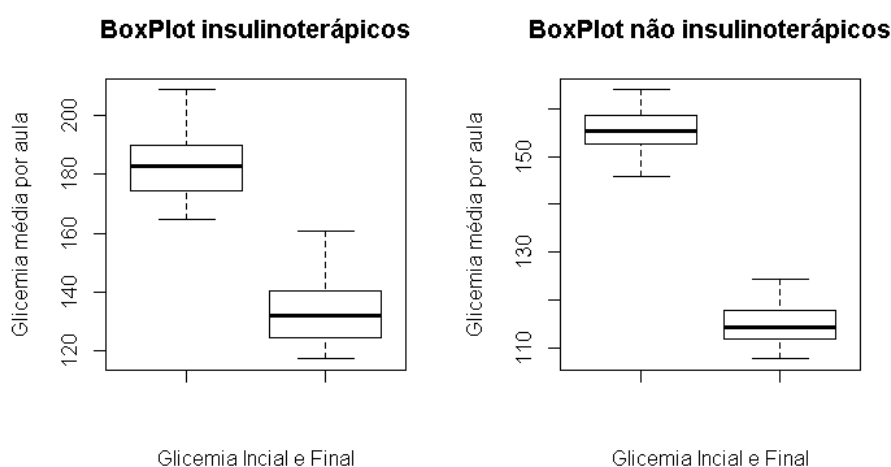


Figura 4.9: Boxplot de glicemia inicial e final dos insulinoterápicos

Observou-se que os insulinoaterápicos iniciam o exercício com taxas de glicemias maiores que os não insulinoaterápicos.

#### 4.1.9 Análise de sexo e glicemia

##### Variação da glicemia

Tabela 4.14: Tabela de sexo e valores de variação da glicemia

Sexo	Variação da Glicemia		
	Mínimo	Média	Máximo
Masculino	-317	-43,65	129
Feminino	-326	-47,38	203

Analisando a variação da glicemia por sexo, pode-se ver que as pessoas do sexo feminino tem uma variabilidade maior, com o valor mínimo menor e o máximo maior quando comparado com os participantes do sexo masculino. Com o intuito de entender melhor quem são os participantes, que apresentaram esses valores extremos, buscou-se as características de cada um, afim de compreender esses valores.

Na tabela 4.14, os valores mínimos se referem a um participante do sexo masculino, diabético TIPO 2, 65 anos de idade e, um participante do sexo feminino, diabético TIPO 1 e com 12 anos de idade, os dois dependentes de insulina exógena. Sendo assim, percebe-se que a única semelhança entre eles é o uso de insulina exógena, que tem uma influência consistente na variação da glicemia, já que o controle é realizado pelo próprio paciente.

Os valores máximos referem-se a um participante do sexo masculino, diabético TIPO 1, 12 anos de idade e, um participante do sexo feminino, diabético TIPO 1 e com 9 anos de idade. Esses dois participantes são dependentes de insulina exógena, o



que remete a mesma conclusão anterior, os dependentes de insulina exógena tendem a ter valores mais extremos de variação da glicemia, por serem eles quem controlam a quantidade e horário da aplicação da reposição da insulina.

Para testar se o sexo dos participantes influencia na variação da glicemia induzida pela atividade física, realizou-se um teste-t para comparar as médias das variações das glicemias por sexo.

Para realizar a comparação, analisou-se primeiramente se as variâncias das variações das glicemias por sexo eram iguais. Para isso foi realizado um teste-F de Fisher para comparar essas variâncias, onde a hipótese nula é que as variâncias dos grupos são iguais.

O p-valor do teste foi igual a 0,0589, maior que o nível de significância de 5%, ou seja, não rejeita-se a hipótese nula. Temos indícios para dizer que não existe diferença significativa entre as variâncias da variação das glicemias.

Sendo assim, o teste mais adequado para comparar as médias das variações de glicemia por sexo é o teste-t para amostras independentes com variâncias iguais.

As hipóteses do teste-t nesse caso são as seguintes:

- $H_0$ : a média das variações das glicemias é igual entre os sexos.
- $H_a$ : a média das variações das glicemias do sexo masculino é menor que a média das variações das glicemias do sexo feminino.

Tabela 4.15: Teste-t de variação da glicemia por sexo

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	-0,6163	0,2697

Após testar as médias das variações das glicemias por sexo, encontrou-se um p-valor de 0,2697, maior que o nível de significância, assim não há indícios para rejeitar a hipótese nula, ou seja, as médias das variações das glicemias são, aproximadamente, as mesmas quando diz respeito ao sexo. A partir dos resultados, pode-se dizer que a média variação da glicemia não é influenciada pelo sexo do participante. O que pode ser evidenciado também no gráfico 4.10.

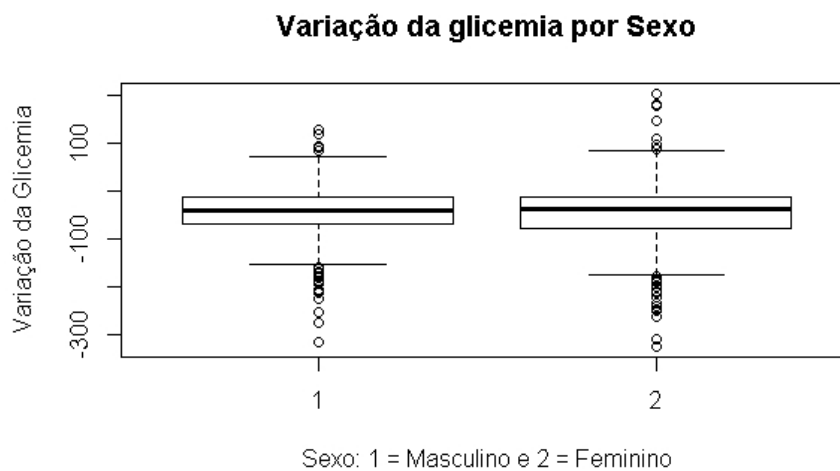


Figura 4.10: Boxplot variação da glicemia por sexo

## Glicemia inicial

Tabela 4.16: Tabela de sexo e glicemias iniciais

Sexo	Valores de glicemias iniciais		
	Mínimo	Média	Máximo
Masculino	48	168,48	466
Feminino	48	167,59	487

Da mesma forma que a análise por variação da glicemia, foram analisados quais os participantes que obtiveram os valores extremos quando diz respeito as glicemias iniciais por sexo, para entender melhor o porquê desses valores.

Na tabela 4.16, para os valores mínimos: três participantes com Diabetes TIPO 1, com idades entre 9 e 21 anos de idade, dois do sexo masculino e um do sexo feminino. Todos apresentaram aumento da glicemia durante as respectivas aulas.

Para os valores de máximo: um participante do sexo masculino, diabético TIPO 2, 65 anos de idade e um participante do sexo feminino, diabético TIPO 1, 9 anos de idade. Os dois apresentaram redução de glicemia no decorrer do exercício. É importante ressaltar que todos os participantes que apresentaram os valores extremos são dependentes de insulina exógena.

## Glicemia final

Tabela 4.17: Tabela de sexo e valores de glicemias finais

Sexo	valores de glicemias finais		
	Mínimo	Média	Máximo
Masculino	18	124,94	430
Feminino	38	120,20	458

A mesma análise foi realizada para os valores de glicemia final, considerando que esses valores foram obtidos após a prática de atividade física.

Na tabela 4.17, para os valores mínimos referem-se a um participante do sexo feminino, diabético TIPO 1, 12 anos de idade e dependente de insulina exógena e, um participante do sexo masculino, diabético TIPO 2, 50 anos de idade e não dependente de insulina exógena. Os valores apresentados foram extremamente baixos, por isso com o acompanhamento dos tutores, esses pacientes possivelmente tenham recebido reposição de glicose, como ,por exemplo, ingerindo algum alimento, aumentar a glicemia.

Para os valores máximos observou-se um participante do sexo masculino, diabético TIPO 2, 65 anos de idade e usuário de insulina exógena e, um participante do sexo feminino, diabético TIPO 1, 48 anos de idade e usuário de insulina exógena. Todos os participantes de valores extremos da tabela 4.17apresentaram redução de glicemia ao longo do exercício físico.

Vale ressaltar que a aula que aparece com maior frequência, em relação às três tabelas de glicemias, é a aula 13.

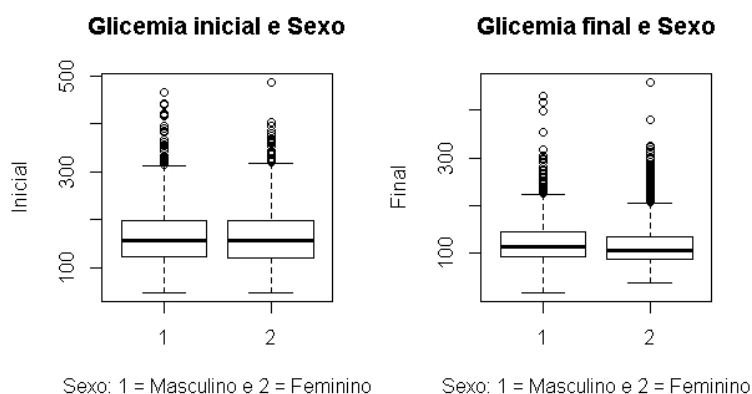


Figura 4.11: Boxplot glicemia inicial e final e sexo

#### 4.1.10 Análise de glicemia por idade

No gráfico de dispersão, figura 4.12, entre idade e variação de glicemia, percebe-se que há maior número de indivíduos com idades avançadas e que as variações são mais concentradas (há uma variabilidade menor) entre as glicemias iniciais e finais dos indivíduos em idades avançadas. A variação da glicemia não está muito correlacionada com a idade dos participantes, apresentando uma correlação de Pearson de 0,16.

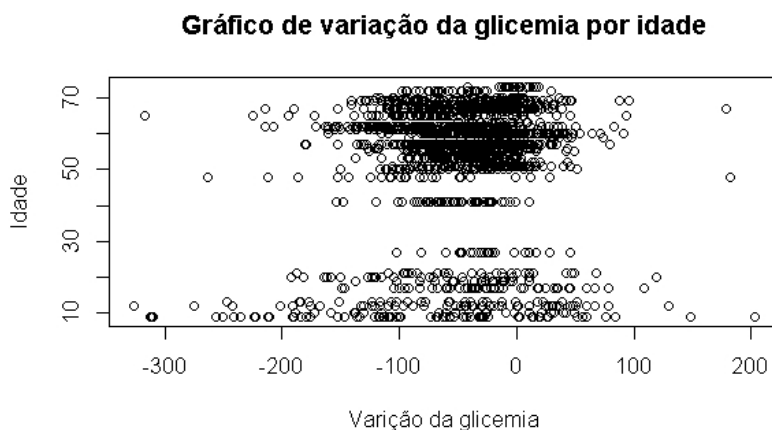


Figura 4.12: Diagrama de dispersão variação da glicemia e idade

#### 4.1.11 Análise de tipo de Diabetes e glicemias

Para analisar se o tipo de Diabetes do participante influencia nas glicemias (inicial, final e variação), foram realizados o teste-t para amostras independentes, com variâncias iguais e tabelas descritivas.

Tabela 4.18: Tabela de tipo de Diabetes e glicemia inicial

Tipo de Diabetes	Glicemia Inicial		
	Mínimo	Média	Máximo
TIPO 1	48	190,14	487
TIPO 2	66	163,71	466

### Glicemia inicial

As hipóteses do teste-t para amostras independentes com variâncias iguais são as seguintes:

- $H_0$ : a média das glicemias iniciais entre os tipos de Diabetes são iguais.
- $H_a$ : a média das glicemias iniciais do Diabetes TIPO 1 é maior que a média das glicemias iniciais do Diabetes TIPO 2.

Tabela 4.19: Teste-t de tipo de Diabetes e glicemia inicial

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	2,2431	0,013

A partir do teste, encontrou-se um p-valor baixo, sendo assim rejeita-se a hipótese nula, o que quer dizer que a média das glicemias iniciais dos diabéticos TIPO 1 é maior que a média das glicemias iniciais dos diabéticos TIPO 2.

## Glicemia final

Tabela 4.20: Tabela de tipo de Diabetes e glicemia final

tipo de Diabetes	Glicemia final		
	Mínimo	Média	Máximo
TIPO 1	38	130,39	458
TIPO 2	18	121,20	430

As hipóteses do teste-t para amostras independentes com variâncias iguais são as seguintes:

- $H_0$ : a média das glicemias finais entre os tipos de Diabetes são iguais.
- $H_a$ : a média das glicemias finais do Diabetes TIPO 1 é maior que a média das glicemias finais do Diabetes TIPO 2.

Tabela 4.21: Teste-t de tipo de Diabetes e glicemia final

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	1,0542	0,1474

De acordo com o teste-t realizado e o p-valor da tabela 4.21, não temos indícios para rejeição da hipótese nula. Não há diferença significativa entre as médias das glicemias finais das aulas com relação ao tipo de Diabetes, diferente do evidenciado com as glicemias iniciais.

## Variação da glicemia

Tabela 4.22: Tabela de tipo de Diabetes e valores de variação da glicemia

Tipo de Diabetes	variação da glicemia		
	Mínimo	Média	Máximo
TIPO 1	-326	-59,74	203
TIPO 2	-317	-42,59	179

As hipóteses do teste-t para amostras independentes com variâncias iguais são as seguintes:

- $H_0$ : a média da amplitude das glicemia entre os tipos de Diabetes são iguais.
- $H_a$ : a média da amplitude das glicemias dos diabéticos TIPO 1 é maior que a média da amplitude das glicemias dos diabéticos TIPO 2.

Tabela 4.23: Teste-t de tipo de Diabetes e variação da glicemia

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	2,23	0,01404

Analisando o p–valor quanto a variação da glicemia também é menor que o nível de significância ( $\alpha$ ), logo, há uma maior variação média significativa no Diabetes TIPO 1 em relação ao Diabetes TIPO 2. Os testes mostram que somente a média da glicemia final não apresentou diferença com relação aos tipos de Diabetes.

#### 4.1.12 Análise de tempo e glicemia

A variável tempo é a soma acumulada dos intervalos de tempo entre as aula, de cada participante.

Tabela 4.24: Correlações de Pearson entre as variáveis

Variáveis	$\rho$
Tempo e glicemia inicial	-0,027
Tempo e glicemia final	-0,025
Tempo e variação da glicemia	0,007

A partir da tabela 4.24 e gráfico 4.13, observa-se a não linearidade da relação entre os dados. Não há influência de uma variável na outra, nem direta ou inversamente



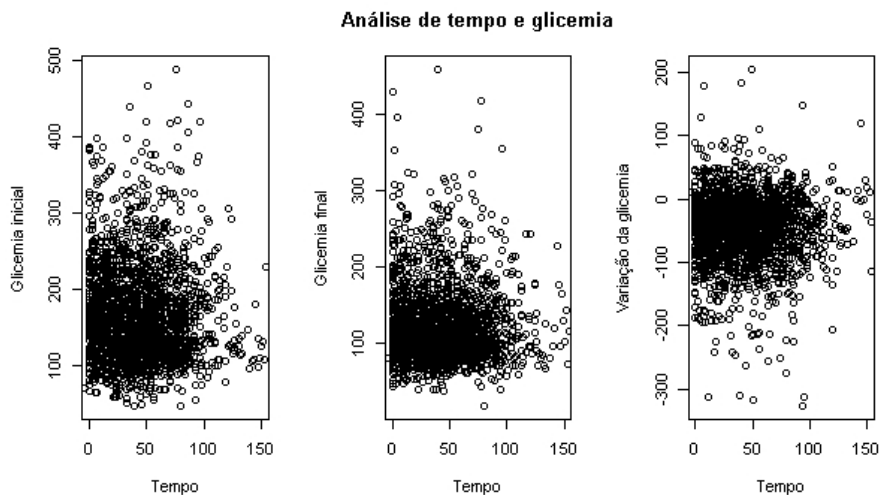


Figura 4.13: Gráficos de dispersão tempo e glicemias

proporcional. As glicemias iniciais, finais e variação da glicemia dos participantes não tem influência do tempo das aulas, a partir da primeira aula, confirmado pelas correlações de Pearson.

## 4.2 Resultados da comparação com relação as aulas

Com o objetivo de mostrar a evolução no controle da glicemia dos participantes através dos exercícios físicos, ao longo das aulas, e a diferença entre duas características em uma mesma aula, realizou-se comparações entre as glicemias finais ou iniciais, sejam elas das primeiras e das últimas aulas de cada participante, levando em consideração uma determinada característica.

Considerou-se, também, as glicemias analisando duas variáveis diferentes em uma determinada aula. Com estas comparações, evidenciou-se algumas diferenças significativas nas glicemias iniciais e finais, confirmadas por testes de hipóteses nas

sessões a seguir.

#### 4.2.1 Comparação entre as aulas por sexo

**Comparação entre glicemias finais, da primeira e última aula, do sexo masculino.**

As hipóteses do teste-t para amostras pareadas são as seguintes:

- $H_0$ : as médias das glicemias finais na primeira e na última aula são iguais em relação aos participantes do sexo masculino.
- $H_a$ : a média das glicemias finais na primeira aula é maior que a média das glicemias finais na última aula, dos participantes do sexo masculino.

Tabela 4.25: Teste-t entre as glicemias finais na 1ª e na última aula para o sexo masculino

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	44	1,9923	0,01329

Os participantes do sexo masculino tiveram uma redução significativa na média das glicemias finais, da primeira para última aula. A aula 1 foi comparada com todas as aulas e, mostrou uma redução significativa, a partir da aula 21. Foi feito o teste, também, para os participantes do sexo feminino mas não se mostrou significativo quanto a diferença entre as médias das aulas.

#### **Comparação das glicemias iniciais, da primeira aula, entre os sexos**

As hipóteses do teste-t para amostras independentes com variâncias diferentes são as seguintes:

- $H_0$ : a média das glicemias iniciais na primeira é igual entre os sexos.

- $H_a$ : a média das glicemias iniciais, do sexo masculino é maior que a média das glicemias iniciais, do sexo feminino, na primeira aula.

Tabela 4.26: Teste-t das glicemias iniciais na 1ª aula entre os sexos

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	1,9589	0,02674

Os indivíduos do sexo masculino iniciavam as atividades físicas com glicemia inicial, significativamente maior, que os indivíduos do sexo feminino.

### 4.2.2 Comparação, entre as aulas, dos usuários de insulina exógena

#### Glicemia inicial na primeira e na última aula

As hipóteses do teste-t para amostras pareadas são as seguintes:

- $H_0$ : a média das glicemias iniciais na primeira e na última aula é igual dos usuários de insulina exógena.
- $H_a$ : a média das glicemias iniciais na primeira aula é maior que a média das glicemias iniciais na última aula dos usuários de insulina exógena.

Tabela 4.27: Teste-t glicemia inicial dos usuários de insulina exógena entre as aulas

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	37	2,3114	0,003083

Os indivíduos insulino-terápicos iniciam as atividades físicas do programa com uma glicemia inicial, significativamente maior, na primeira aula em relação à última aula.

## Comparação das glicemias finais, na primeira e na última aula

As hipóteses do teste-t para amostras pareadas são as seguintes:

- $H_0$ : a média das glicemias finais na primeira e na última aula é igual entre os usuários de insulina exógena.
- $H_a$ : a média das glicemias finais, dos usuários de insulina exógena é maior na primeira aula do que a média das glicemias finais, dos usuários de insulina exógena, na última aula.

Tabela 4.28: Teste-t glicemia final entre aulas

	$\alpha$	Graus de liberdade	T	p-valor
Teste T	0,05	37	2,9049	0,0008843

Os dependentes de insulina exógena tiveram uma diferença na glicemia final da sua primeira aula para última, especificamente, a partir da sexta aula, a partir da comparação da primeira aula com as seguintes.

### 4.2.3 Comparação das glicemias iniciais e finais das aulas, dos usuários e não usuários de insulina exógena

#### Comparação das glicemias iniciais, na primeira aula

As hipóteses do teste-t para amostras independentes com variâncias diferentes, são as seguintes:

- $H_0$ : a média das glicemias iniciais na primeira aula é igual entre os usuários, e não usuários de insulina exógena.

- $H_a$ : a média das glicemias iniciais, dos usuários de insulina exógena, é maior que a média das glicemias iniciais, dos não usuários de insulina exógena, na primeira aula.

Tabela 4.29: Teste-t glicemia inicial primeira aula

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	-4,0844	0,00005

Com o resultado do teste-t, pode-se afirmar que os usuários de insulina exógena iniciavam a atividade física com média de glicemia maior que os não usuários de insulina exógena.

### Comparação das glicemias finais, na primeira aula

As hipóteses do teste-t para amostras independentes com variâncias diferentes são as seguintes:

- $H_0$ : a média das glicemias finais na primeira aula é igual entre os usuários, e não usuários, de insulina exógena.
- $H_a$ : a média das glicemias finais, dos não usuários de insulina exógena, é menor que a média das glicemias iniciais, dos não usuários de insulina exógena, na primeira aula.

Tabela 4.30: Teste-t glicemia final primeira aula

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	-3,8903	0,0001

Com um p-valor abaixo do nível de significância  $\alpha$ , opta-se pela rejeição da hipótese nula, onde a média das glicemias finais dos usuários de insulina exógena é maior que a média das glicemias finais dos não usuários.

## Comparação das glicemias finais, na última aula

As hipóteses do teste-t para amostras independentes com variâncias diferentes são as seguintes:

- $H_0$ : a média das glicemias finais na última aula é igual entre os usuários, e não usuários de insulina exógena.
- $H_a$ : a média das glicemias finais, dos usuários de insulina exógena, é maior que a média das glicemias finais, dos não usuários de insulina exógena, na última aula.

Tabela 4.31: Teste-t glicemia final última aula

	$\alpha$	Graus de liberdade	T	p-valor
Teste-t	0,05	83	-0,0495	0,4839

Dos usuários de insulina exógena e não usuários, na primeira aula havia diferença entre as glicemias finais, mas na última aula não houve. Antes os insulino-terápicos terminavam com uma glicemia final mais alta que os não usuários. A partir da 13<sup>a</sup> aula a média das glicemias finais daqueles que dependem, e daqueles que não dependem, de insulina exógena já não apresentaram diferença significativa, na comparação da primeira aula com as seguintes.

#### 4.2.4 Comparação das médias iniciais e finais, por aula

Foram calculadas as médias, de todos os participantes, das glicemias iniciais e finais, por aula, e colocadas no gráfico 4.14 para melhor visualização da variação entre as elas.

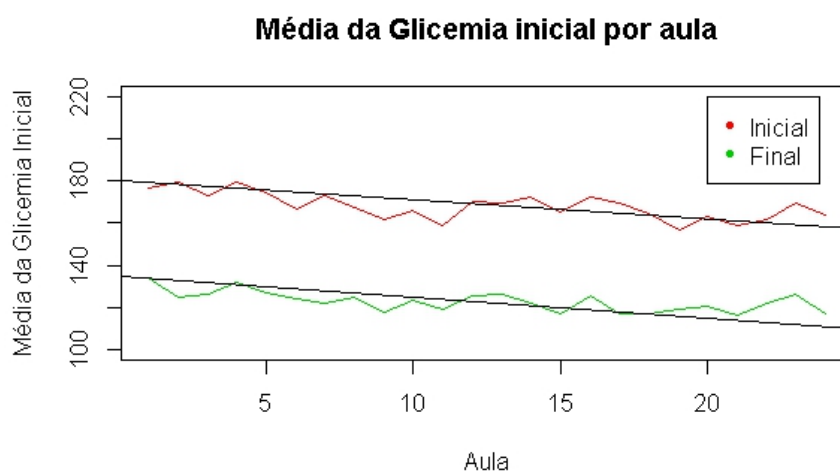


Figura 4.14: Médias das glicemias iniciais e finais por aula

O gráfico 4.14 confirma os resultados apresentados anteriormente, pois há uma nítida diferença entre as médias das glicemias finais e iniciais em todas as aulas.

# Capítulo 5

## Modelo dos dados

### 5.1 Ajuste do modelo

Encontrar o modelo ideal para certos dados significa encontrar as variáveis do banco de dados que descrevam o comportamento da variável de interesse (variável resposta). A partir das variações das glicemias coletadas de cada participante do Programa "Doce DESAFIO", foram usados, como dito anteriormente, modelos lineares mistos. Pela análise descritiva e pelo gráfico 5.1, percebe-se que as curvas das variações glicêmicas dos participantes se iniciam em pontos diferentes, sugerindo assim a inserção no modelo de um intercepto aleatório.

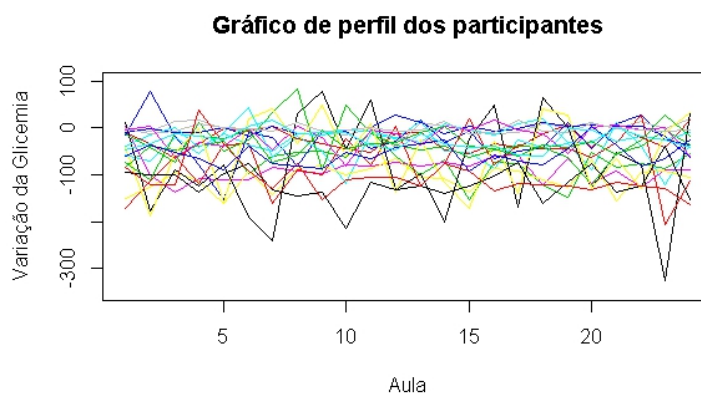


Figura 5.1: Gráfico de perfil dos participantes



O gráfico 5.1 apresenta dados de apenas 20 dos 85 participantes, selecionados aleatoriamente, para garantir uma melhor visualização.

Definida a presença de um intercepto aleatório, necessitou-se do cálculo dos demais efeitos fixos para completar o modelo. Estimou-se estes efeitos pelo método de Máxima Verossimilhança e, foram retiradas as variáveis menos significativas para o modelo, uma a uma. Após aproximar o modelo ideal aos dados, estimou-se este modelo por Máxima Verossimilhança Restrita aplicada a duas estruturas de covariância. Não foi aplicado à estrutura de covariância Não Estruturada por dificuldades computacionais.

### 5.1.1 Modelos estimados

Tabela 5.1: Modelo 1 (g1) : variação da glicemia  $\sim$  tempo + sexo + insulino-terapia + idade + tipo de Diabetes

	Value	Std,Error	DF	t-value	p-value
(Intercept)	-43,19812	5,440082	2011	-7,94071	0
tempo	0,00597	0,01786	2011	0,334077	0,7384
Sexo2	-4,26022	2,393528	2011	-1,77989	0,0752
Insul1	-7,46489	2,785221	2011	-2,680178	0,0074
Idade	-0,3767	0,141825	2011	-2,656108	0,008
TipoDM2	27,34766	6,668614	2011	4,100951	0

Tabela 5.2: Modelo 2 (g2) : variação da glicemia  $\sim$  insulino-terapia + sexo + idade + tipo de Diabetes

	Value	Std,Error	DF	t-value	p-value
(Intercept)	-42,93428	5,381277	2012	-7,978456	0
Insul1	-7,46744	2,784603	2012	-2,681688	0,0074
Sexo2	-4,27673	2,392495	2012	-1,78756	0,074
Idade	-0,37579	0,141768	2012	-2,650758	0,0081
TipoDM2	27,30853	6,66613	2012	4,09661	0

Tabela 5.3: Modelo 3 (g3) : variação da glicemia  $\sim$  insulinoaterapia + idade + tipo de Diabetes

	Value	Std,Error	DF	t-value	p-value
(Intercept)	-46,86849	4,913272	2013	-9,539162	0
Insul1	-6,14406	2,685834	2013	-2,287582	0,0223
Idade	-0,3667	0,141753	2013	-2,586861	0,0098
TipoDM2	28,32293	6,645512	2013	4,261964	0

Tabela 5.4: Modelo 4 (g4) : variação da glicemia  $\sim$  idade + tipo de Diabetes

	Value	Std,Error	DF	t-value	p-value
(Intercept)	-53,93474	3,824774	2014	-14,10142	0
Idade	-0,31646	0,140187	2014	-2,257409	0,0241
TipoDM2	30,30312	6,595737	2014	4,59435	0

Tabela 5.5: Modelo 5 (g5) : variação da glicemia  $\sim$  tipo de Diabetes

	Value	Std,Error	DF	t-value	p-value
(Intercept)	-59,74405	2,832385	2015	-21,093197	0
TipoDM2	17,15602	3,099078	2015	5,535846	0

Analisando os modelos, verificou-se quais variáveis estavam abaixo no nível de significância  $\alpha$  igual a 5%. Retirando as variáveis uma a uma, quanto maior o p-valor menos significativa é essa variável para o modelo.

O modelo g1 é o modelo com todas as variáveis presentes no banco de dados. As variáveis de g1 foram: tempo, sexo, insulinoaterapia, idade e tipo de Diabetes. Em g1, a variável menos significativa foi "tempo", pois apresentou um p-valor maior que o  $\alpha$  logo, ela foi retirada do modelo.

As variáveis do próximo modelo g2, sem "tempo", foram: insulinoaterapia, sexo, idade e tipo de Diabetes. Em g2, a variável que se apresentou menos significativa foi "sexo" e, assim, e foi retirada do modelo.

As variáveis do próximo modelo g3 foram: insulinoaterapia, tipo de Diabetes

e idade. Em g3, todas as variáveis apresentaram o p–valor menor que o nível de significância, logo todas são significativas para o modelo. Mesmo sendo todas significativas, foi retirada a variável menos significativa, ou seja, a com maior p–valor que foi "insulinoterapia".

No novo modelo, g4, as variáveis que seguiram no modelo foram: idade e tipo de Diabetes. As duas foram significativas para o modelo, ou seja, as variáveis idade e tipo de Diabetes são explicativas do modelo, que remetem a variável resposta de interesse variação da glicemia.

Mais uma vez, apesar de todas as variáveis se apresentarem significativas ao modelo, retira-se "idade". E o modelo g5, o mais simples, ficou apenas com tipo de Diabetes como variável explicativa.

### **ANOVA de comparação entre os modelos com as variáveis significativas**

Após a seleção das variáveis explicativas mais significativas, comparou-se o modelo mais completo de variáveis significativas (g3) com os mais simples (g4 e g5):

Tabela 5.6: Tabela de comparação entre os modelos estimados pelo método MV

Model	df	AIC	BIC	logLik	Comparação	L.Ratio	p-value
g3	6	21903,61	21937,33	-10945,8			
g4	5	21906,84	21934,95	-10948,42	g3 x g4	5,236584	0,0221
g3	6	21903,61	21937,33	-10945,8			
g5	4	21909,94	21932,42	-10950,97	g3 x g5	10,33361	0,0057
g4	5	21906,84	21934,95	-10948,42			
g5	4	21909,94	21932,42	-10950,97	g4 x g5	5,097028	0,024

Pelo teste de razão de verossimilhança, o modelo mais completo, g3, foi o melhor dentre os estimados.

## Comparação entre estruturas de covariância aplicadas ao modelo estimado

Na estimação do modelo g3, estimado por MVR, aplicado a duas diferentes estruturas de covariância (VC e CS), obteve-se:

Tabela 5.7: Tabela de comparação entre os métodos de estimação dos modelos

Modelo	Matriz	AIC	BIC
g3	VC	21895,68	21929,39
g3.1	CS	21894,67	21934,00

Optou-se pelo modelo g3 utilizando a matriz VC, devido ao seu menor valor de BIC, melhor indicado para decisão quando há divergência entre os modelos.

## Gráficos de resíduos

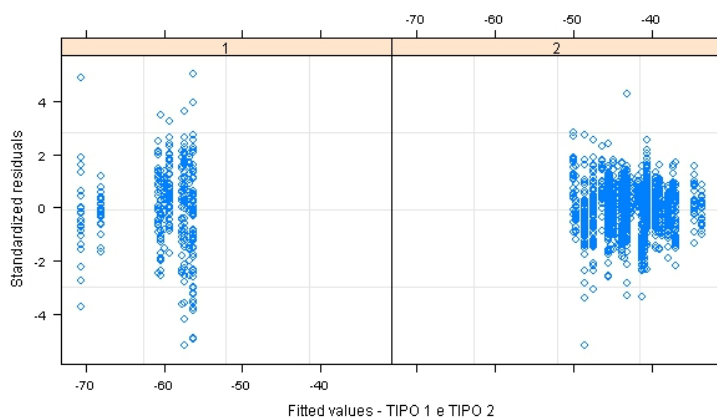


Figura 5.2: Gráfico dos resíduos padronizados

No gráfico 5.2 foi gerado da padronização dos resíduos de g3 por tipo de Diabetes, a variável mais significativa. A forma pela qual os pontos estão dispersos pelo gráfico não é aleatória, ou seja, não há a indicação de aleatoriedade nos resíduos do modelo.

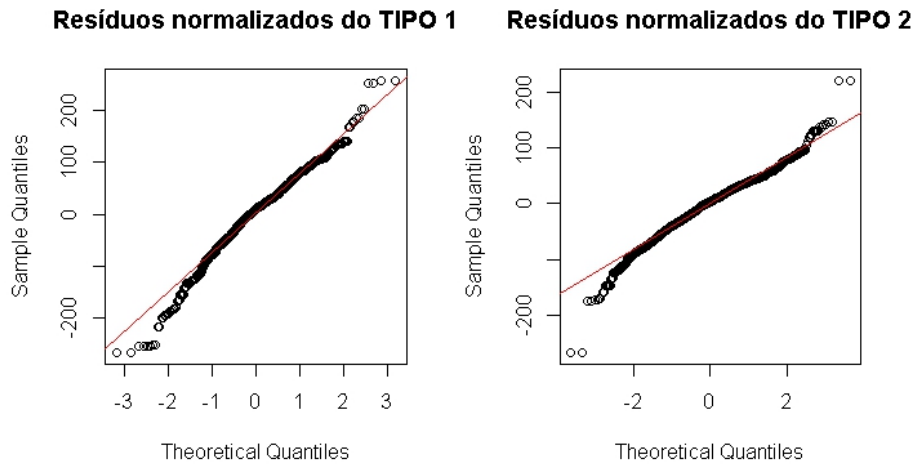


Figura 5.3: Gráfico dos resíduos normalizados

No gráfico 5.3, da normalização dos resíduos, verificou-se que os erros apresentaram distribuição aproximada da Normal, como pode-se notar analisando a linha de tendência do gráfico que contempla a maioria dos pontos. Tendo em vista as conclusões com relação a não aleatoriedade dos resíduos, não se percebeu um ajuste adequado do modelo aos dados do "Doce DESAFIO".

# Capítulo 6

## Conclusão

A partir de todas as análises descritivas realizadas neste trabalho, concluiu-se que há a mesma probabilidade de incidência do Diabetes TIPO 1 e TIPO 2 em ambos os sexos. Há, também, a mesma probabilidade da necessidade do uso de insulina exógena em ambos os sexos e, não foi evidenciada diferença significativa das idades entre os sexos. Foi evidenciada diferença significativa das idades, tanto por insulino terapia, quanto por tipo de Diabetes. Os usuários de insulina exógena e diabéticos TIPO 1 apresentaram idades menores.

No caso das correlações de Pearson das variáveis quantitativas, idade e variação da glicemia, não houve uma correlação significativa, ou seja, a idade não influencia significativamente na variação da glicemia. Chegou-se na mesma conclusão na correlação entre tempo e variação da glicemia.

Na comparação de glicemias iniciais e glicemias finais, houve diferença significativa, confirmando a eficácia do exercício físico na redução da glicemia inicial. Os diabéticos TIPO 1 possuem maiores glicemias iniciais e maiores variações das glicemias do que os diabéticos TIPO 2, o que pode ser influenciado pela necessidade de insulina exógena.

Partindo para resultados em relação aos efeitos das aulas, os participantes do sexo masculino tiveram uma redução de suas glicemias finais, da primeira para a última aula do programa. O mesmo para os usuários de insulina exógena. O usuários de insulina exógena mantiveram glicemias iniciais maiores que os não usuários, da primeira para a última aula, mas em relação as glicemias finais, que eram maiores na primeira aula, se tornaram iguais, na última aula.

O modelo final ficou composto pelas as variáveis explicativas idade, tipo de Diabetes e insulino terapia, estas variáveis são as que explicam significativamente a variação da glicemia dos participantes. Vale ressaltar que um dos pressupostos não foi atendido, aleatoriedade dos resíduos, mas para os demais o modelo se mostrou válido.

# Referências

- BARBOSA, M. **Uma abordagem para análise de dados com medidas repetidas utilizando modelos lineares mistos**. 2009. Tese (Mestrado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba 2009.
- CAITANO P.; WINKELMANN E.; BERLEZI E. **Perfil de indivíduos com Diabetes Mellitus TIPO 2 adscritos em estratégia de saúde da família**
- CROWDER, M.J.; HAND, D. **Analysis of repeated measures**. 257p.,1990.
- DIGGLE, P. **An approach to the analysis of repeated measurements**. 1988, p. 959-971.
- DULLIUS, J. **Diabetes Mellitus - Saúde, educação, atividade físicas**. 2007, 418p.
- LAIRD, N.M.; WARE ,J.H. **Random effects models for longitudinal data**. Biometrics, Washington, 1982.
- LIMA, C.G. **Análise de dados longitudinais provenientes de experimentos em blocos casualizados**. 1996 119 p. Tese (Doutorado em Estatística e Experimentação Agronômica) Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1996.



PATTERSON, H.D.; THOMPSON, R. **Recovery of interblock information when block sizes are unequal**. 1971, p. 545-554.

PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer - Verlag, 2000, 528p.

VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models for longitudinal data**. New York: Springer - Verlag, 2000, 568p.

ZIMONEM B.; RUDERMAN N., **Posicionamento Oficial Diabetes Mellitus e Exercício**, 1997.